

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 06:04:49 ; Search time 52 Seconds
(without alignments)
458.880 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCPLAPGHVSQALGTLT.....PLGALLELLSQPLFPYADP 248
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103.5	7.9	550	2 S64725	probable lipoprote
2	94.5	7.2	393	2 T07167	probable isocitrat
3	90.5	6.9	271	2 T36895	probable membrane
4	90	6.9	1088	2 H96747	unknown protein T1
5	89	6.8	798	2 A12053	competence protein
6	86.5	6.6	929	2 A44048	genome polypeptide
7	86	6.6	493	1 S51574	mocr protein - Rhi
8	85.5	6.5	420	2 A82699	conserved hypotet
9	85.5	6.5	468	1 B46114	glycoprotein gp13
10	85.5	6.5	2303	1 GNNVTP	genome polypeptide
11	85.5	6.5	2303	2 S13554	genome polypeptide
12	84.5	6.5	289	2 A83608	polyamine transpor
13	83.5	6.4	468	1 VGBEEH	glycoprotein gp13
14	83	6.3	550	1 VGBE18	glycoprotein E - h
15	83	6.3	4436	2 T71086	hypothetical prote
16	82.5	6.3	233	2 T15620	hypothetical prote
17	82.5	6.3	288	2 A55737	PD-1 protein - hum
18	82.5	6.3	636	2 T72736	hypothetical prote
19	82	6.3	398	2 T46475	hypothetical prote
20	81	6.2	450	1 FOLJFP	gag polypeptide -
21	81	6.2	450	2 S23819	gag protein - feli
22	80.5	6.2	312	2 C87562	conserved hypotet
23	80.5	6.2	541	2 T48811	hypothetical prote
24	80.5	6.2	1240	2 T03097	CDO protein - huma
25	80.5	6.2	2301	1 GNNVTM	genome polypeptide
26	80.5	6.2	2303	1 GNNVTM	genome polypeptide
27	80	6.1	290	2 T09260	aquaporin-like tra
28	79	6.0	407	2 B70962	hypothetical prote
29	79	6.0	950	2 T51134	ionotropic glutama

30	78	6.0	204	2 T51012	hypothetical prote
31	77.5	5.9	390	1 Q0BE77	glycoprotein I pre
32	77.5	5.9	834	2 S66498	M-sena F protein p
33	77	5.9	99	2 C46518	Ig L1 chain V regi
34	77	5.9	124	2 S08640	hypothetical prote
35	77	5.9	287	2 S60455	transmembrane prot
36	77	5.9	289	2 S33617	trg-31 protein - g
37	77	5.9	303	2 A40807	membrane glycoprot
38	77	5.9	364	2 H75466	lipopolysaccharide
39	77	5.9	442	1 XUBRVS	3-phosphoshikimate
40	77	5.9	910	2 D75524	alpha-dextran endo
41	77	5.9	3097	2 T00021	DN-cadherin - frui
42	77	5.9	26926	1 I38344	titin, cardiac mus
43	76.5	5.8	351	2 S39603	class I histocompa
44	76.5	5.8	666	2 H83943	transketolase tkt
45	76.5	5.8	835	2 S40140	ribonucleoside-dip

ALIGNMENTS

RESULT 1
S64725
probable lipoprotein uxpa precursor - Pseudomonas putida
C;Species: Pseudomonas putida
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S64725; S47503
R;de Groot, A.; Kriliger, J. J.; Filloux, A.; Tommassen, J.
Mol. Gen. Genet. 250, 491-504, 1996
A;Title: Characterization of type II protein secretion (xcp) genes in the plant growth-
A;Reference number: S64724; MUID:96186881; PMID:8602167
A;Accession: S64725
A;Molecule type: DNA
A;Residues: 1-550 <DEG>
A;Cross-references: UNIPROT:Q52289; EMBL:X81085; NID:g3293032; PIDN:CAA56977.1; PID:g53
C;Genetics:
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-550/Product: probable lipoprotein uxpa #status predicted <NAT>

Query Match	7.9%	Score 103.5;	DB 2;	Length 550;
Best Local Similarity	23.2%	Pred. No. 0.17;		
Matches	72;	Conservative 42;	Mismatches 118;	Indels 79;
Gaps	16;			
QY	2	QTCPLAPGHVSQA-LGTLFLAA-----SLSAQN-EGWDSPTCTGVSVVS	46	
Db	90	QLAPLLDAGRADEANLGGVAVLGGALLEQLRGEAGAGNSLTLENGQWNG----SGLAYLT	145	
QY	47	WGVNTVMSNCINAFSHVNIKLR--HGQESAIFFNEVAPGYFSR-----DGWQ	92	
Db	146	QGESGVQSLLGSEARVSSDERVLPQRSAAALYRQASATTLGAGLADEQRQALGLEPLQ	205	
QY	93	LQVQGVQAOLVTKG-----ARDSHAGLYMWH---LVGHORNNRQVTLVSGASPSQAP--	142	
Db	206	LPERGS-ARIAVGVVTDTPYAQDKASLKQYQALLFPVFOQARREADLVVALADVGTGPG	264	
QY	143	-----DTGFVPVPVAVTAVFILLVALVMPAW-----YRCRC-----SQQ	176	
Db	265	WLAERLPAIDLLLCARGQDLWPTPVQATQASGRVP-VLFAGCRGSGAFELRCQQVAGQW	323	
QY	177	REKKFF-----LLEPQMKVAALRAGAQQGLSRASAEIWTDPDSEPTPR-PLALVFPSP	230	
Db	324	QFEGRFFPTFEQTLSPAQVRAGQLOAQNLQORAGHAWL--DOPLARAPQALWRRDRTRG	381	
QY	231	GALELLSQPL 241		
Db	382	GSWDRLLHQA 392		

RESULT 2
T07167
Probable isocitrate dehydrogenase (NAD) (EC 1.1.1.41) chain 1 - tomato
C;Species: Lycopersicon esculentum (tomato)

[illegible]

F;1065-1191/Product: protein 2B (core protein P2-5b) #status predicted <P2B> F;1192-1511/Product: protein 2C (core protein P2-X) #status predicted <P2C> F;1518-1605/Product: protein 3A (protein P3-1b) #status predicted <P3A> F;1606-1625/Product: protein 3B (genome-linked protein VPg) #status predicted <P3B> F;1626-1842/Product: protein 3C (probable proteinase) #status predicted <P3C> F;1843-2303/Product: protein 3D (probable RNA-directed RNA polymerase) #status predicted
Query Match Best Local Similarity 6.5%; Score 85.5; DB 1; Length 2303; Matches 28; Conservative 16; Mismatches 33; Indels 25; Gaps 6;
Qy 60 AFSHVNIKLRAH---GOESAIFN-EVAPGYPSRDLGWLQGVGGVAQLVIKGARDSHAGLY 115 : : : : Db 223 AFSHIRIPL-PHVLAGEDGGVFGLTARRHYLCKTGMRVQVCNASQF-----HAGSL 273 : : : : Qy 116 MWHLV-----GHQRNNRVITLEVSQAEPQSADPTGCF 146 : : : : Db 274 LVFMWAFYFTGKGTSGMTSPDFDTMDTTWRSPQSAP-TGY 314 : : : :
RESULT 11 S13554 Genome polyprotein - murine poliovirus C:Species: murine poliovirus, Theiler's encephalomyelitis virus C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004 C:Accession: S13554 R;Law, K.M.; Brown, T.D.K. Nucleic Acids Res. 18, 6707-6708, 1990 A>Title: The complete nucleotide sequence of the GPVII strain of Theiler's murine encephalomyelitis virus A:Reference number: S13554; MUID: 91067481; PMID: 2251141 A:Accession: S13554 A>Status: preliminary; translation not shown A:Molecule type: mRNA A:Residues: 1-2303 <LAW> A:Cross-references: UNIPROT:Q88595; EMBL:X56019; NID:G62039; PIDN:CAA39496.1; PID:G62040 C:Superfamily: foot-and-mouth disease virus genome polyprotein
Query Match Best Local Similarity 6.5%; Score 85.5; DB 2; Length 2303; Matches 28; Conservative 16; Mismatches 33; Indels 25; Gaps 6;
Qy 60 AFSHVNIKLRAH---GOESAIFN-EVAPGYPSRDLGWLQGVGGVAQLVIKGARDSHAGLY 115 : : : : Db 223 AFSHIRIPL-PHVLAGEDGGVFGLTARRHYLCKTGMRVQVCNASQF-----HAGSL 273 : : : : Qy 116 MWHLV-----GHQRNNRVITLEVSQAEPQSADPTGCF 146 : : : : Db 274 LVFMWAFYFTGKGTSGMTSPDFDTMDTTWRSPQSAP-TGY 314 : : : :
RESULT 12 AB3608 polylamine transport protein potI PA0304 [imported] - Pseudomonas aeruginosa (strain PAC) C:Species: Pseudomonas aeruginosa C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004 C:Accession: AB3608 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Eickholt, J.; Lau, B.; Olson, M.V.; Adams, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lane, D.C.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen A:Reference number: AB3608; MUID: 20437337; PMID: 10984043 A:Accession: AB3608 A>Status: preliminary A:Molecule type: DNA A:Residues: 1-289 <STO> A:Cross-references: UNIPROT:Q91617; GB:AE004468; GB:AE004091; NID:g9946144; PIDN:AAG036 A:Experimental source: strain PAO1 C:Genetics: A:Gene: potI; PA0304 C:Superfamily: spermidine/putrescine transport system permease protein potI
Query Match Best Local Similarity 6.5%; Score 84.5; DB 2; Length 289;


```

QY      2  QTCPLAFPHGVHQSALGTLTLFLAASLSAQNECHDSP-----ICTEGVVSVSXGENTVMS 54
DB      34  QSTP-ATPHTPNLTAAHAGASDNTNANGTESTHSHETTTCTKSLISVPYKSDMN 92
QY      55  CNISNAFSHVNIKLRAHGOESAIFNEVAPG---YF-----SRDG 90
DB      93  CTTSGVGVNYSYRLEIYLNORTPFGCTPGDEENYLNNHNAKDQTLFLFSTAERKKSRRG 152
QY      91  WQLQV-----OGGVA-OLVTIKGARDSHAGLYMMHLV---GHQRNRRQV 129
DB      153  GOLGVIPDLPRKQFLNPLHTEGGFKFPLTIKSDVWRTAGIYVMSLYAKNGTLVNSTSV 212
QY      130  TLEVSQA 136
DB      213  TVSTYNA 219

RESULT 14
VGBE18
glycoprotein E - human herpesvirus 1
N:Alternate names: US8
C:Species: human herpesvirus 1
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A03733; A45696
R:MGGocho, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.
J. Mol. Biol. 181, 1-13, 1985
A:Title: Sequence determination and genetic content of the short unique region :
A:Reference number: A00656; MUID:85160822; PMID:2984429
A:Accession: A03733
A:Molecule type: DNA
A:Residues: 1-550 <MCG>
A:Cross-references: UNIPROT:P04488; GB:X02138; NID:G59865; PIDN:CAA26062.1; PID
A:Experimental source: strain 17
R:Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.
J. Virol. 67, 3961-3968, 1993
A:Title: Identification of a new transcriptional unit that yields a gene product
A:Reference number: A45696; MUID:93287213; PMID:8389914
A:Accession: A45696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 438-550 <GEO>
A:Cross-references: GB:S62895; NID:G386127; PIDN:AAB27080.1; PID:G386128
A:Experimental source: R35
A:Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBI:P133647)
C:Superfamily: herpesvirus glycoprotein E
C:Keywords: glycoprotein
F:124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.3%; Score 83; DB 1; Length 550;
Best Local Similarity 25.4%; Pred. No.12;
Matches 53; Conservative 16; Mismatches 62; Indels 78; Gaps 11;

QY      80  EVAPGVFSRDGWLQVGGVAQLVIKGARDSHAGLYMW-----HLVGH-----QR 124
DB      330  EFPVG-----LAWQAASVNLFRDASPSQSHGLYLCVVYVNDHIHANGHITITAAQY 381
QY      125  NNRQV-----TLEVGAEPQASPDTEGFWPVPVAVTAVFILLVALVWFAYWR 170
DB      382  RRAVVEQLPQRGADLAETPHVGVAPPHPATHGALRLGAVMGAA-LLLSALGLSVWAC 440
QY      171  CRCQORREKKFFLEPQMKVAALRAGAQQGLSR-----ASAEL---WTPDSE----- 215
DB      441  MTCWRR-----AWRAVKSRAQKGFYIRVADSELYADWSSDSEGERDQ 485
QY      216  -----PTPRPLALVPKPSPLGA-LELLSP 238
DB      486  VPLWAPPERP-----DSPSTNGSGFETLSP 510

RESULT 15
E71086
hypothetical protein PH0954 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii

```


C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
 C;Accession: E71086
 R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A;Reference number: A71000; MUID:98344137; PMID:9679194
 A;Accession: E71086
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4436 <KAW>
 A;Cross-references: UNIPROT:O58559; GB:AP000004; NID:G3236131; PIDN:BAA30051.1; PID:d103
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C;Genetics:
 A;Gene: PH0954

Query Match 6.3%; Score 83; DB 2; Length 4436;
 Best Local Similarity 23.0%; Pred. No. 1.4e+02;
 Matches 37; Conservative 24; Mismatches 58; Indels 42; Gaps 8;
 Qy 34 DSPICTEGVSVS-----WGENTVMSCNISNMFSHVNIKLR-----HGQE 74
 Db 300 DDPDLNNLAEGVWPGDYWENASINNLIPEFASINFKVTTTSKIPSAKVLLRNGVE 359
 Qy 75 SAI-----FNEVAPGYFSRDGMQLQVQGG--VAQLVIKGARDSHA-----GLYMMHLV 120
 Db 360 EKIEYLSFYNGIAEGEIS--W--LVQGGNYTLALLVEGKGIDINSNNIYLLGNYNPLP 414
 Qy 121 GHQRNROVTL-----EVSQAEPSQAPDTGFWPVPVAVTVAF 157
 Db 415 NFEVGNYSIDLPTCVDSTGEVRVNVVTSTANWSIPVRLTLVY 455

Search completed: October 17, 2005, 07:13:18
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 06:04:49 ; Search time 121 Seconds
(without alignments)
1049.551 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MQTCPLAPGHVSQALGTLT.....PLGALELLSQPLFPYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1308	100.0	248	1 SCTM_HUMAN
2	382	29.2	192	Q211w8 mus musculus
3	376.5	28.8	201	Q6P781
4	358	27.4	201	Q6AYS0
5	327.5	25.0	211	Q8VDU7
6	326	24.9	212	1 SCTM_MOUSE
7	325	24.8	212	Q8D966
8	324	24.8	212	Q8CE38
9	103.5	7.9	550	Q52289
10	102.5	7.8	506	1 SHS1_BOVIN
11	101	7.7	334	Q8XZ39
12	98.5	7.5	304	Q8BPN5
13	98	7.5	542	2 Q8NHN5
14	97	7.4	894	2 Q8V5F7
15	96.5	7.4	550	2 Q8BP09
16	96	7.3	442	2 Q8H8B3
17	96	7.3	917	2 Q6NNX3
18	96	7.3	969	2 Q86K66
19	95	7.3	1140	2 Q80T91
20	94.5	7.2	393	2 Q82004
21	94.5	7.2	690	2 Q62AS3
22	94.5	7.2	690	2 Q63JU4
23	91.5	7.0	864	2 Q6P779
24	91.5	7.0	1260	2 Q7NR05
25	91	7.0	6620	2 Q96AA2
26	90.5	6.9	271	2 Q9S218
27	90	6.9	1088	2 Q9C9D7
28	89	6.8	798	2 Q8YVJ4
29	88.5	6.8	352	2 Q63NY8
30	88	6.7	327	2 Q8RYG9
31	87.5	6.7	298	2 Q804R4

32	87.5	6.7	372	2	Q90Y50
33	87.5	6.7	645	2	Q8MJZ5
34	87.5	6.7	947	2	Q8BKK7
35	87	6.7	259	2	Q82SL8
36	86.5	6.6	395	2	Q88LS8
37	86.5	6.6	420	2	Q87DX4
38	86.5	6.6	930	2	Q02472
39	86	6.6	493	1	MOCR_RHIME
40	86	6.6	508	2	Q80ZD5
41	85.5	6.5	407	2	Q9D2J4
42	85.5	6.5	418	2	Q8MI84
43	85.5	6.5	420	2	Q9PD55
44	85.5	6.5	430	2	Q8MI85
45	85.5	6.5	922	2	Q88495

ALIGNMENTS

RESULT 1

ID	SCTM_HUMAN	STANDARD;	PRT;	248 AA.
AC	Q8WVJ6; Q00466;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Secreted and transmembrane protein 1 precursor (Protein K12).			
GN	Name=SECTM1;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.			
RX	MEDLINE=98149980; PubMed=9480746; DOI=10.1006/geno.1997.5151;			
RA	Slentz-Kessler K.A., Hale L.P., Kaufman R.E.;			
RT	"Identification and characterization of K12 (SECTM1), a novel human			
RT	gene that encodes a Golgi-associated protein with transmembrane and			
RT	secreted isoforms."			
RL	Genomics 47:327-340(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Liver;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Stausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[3]			
RP	INTERACTION WITH CD7.			
RX	MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;			
RA	Lyman S.D., Escobar S., Rouseau A.-M., Armstrong A., Fanslow W.C.;			
RT	"Identification of CD7 as a cognate of the human K12 (SECTM1)			
RT	protein."			
RL	J. Biol. Chem. 275:3431-3437(2000).			
CC	-1- SUBUNIT: Interacts with CD7.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also			

CC found as secreted.
CC TISSUE SPECIFICITY: Detected at the highest levels in peripheral
CC blood leukocytes and breast cancer cell lines. Found in leukocytes
CC of the myeloid lineage, with the strongest expression observed in
CC granulocytes and no detectable expression in lymphocytes.
CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U77643; AAC52044.1; -;
DR EMBL; BC017716; AAH17716.1; -;
DR Genew; HGNC:10707; SECTM1.
DR H-InvDB; HIX0014254; -;
DR MIM; 602602; -;
DR InterPro; IPR007110; Ig-like.
DR SIGNAL; Transmembrane.
KW SIGNAL 1 28

FT CHAIN 29 248 Potential.
FT DOMAIN 29 145 Secreted and transmembrane protein 1.
FT TRANSMEM 146 166 Extracellular (Potential).
FT DOMAIN 167 248 Potential.
FT DISULFID 37 55 Cytoplasmic (Potential).
FT CARBOHYD 56 56 Potential.
FT CONFLICT 191 191 N-linked (GlcNAc...) (Potential).
FT V -> F (in Ref. 2).
SQ SEQUENCE 248 AA; 27039 MW; 21E3066B67920487 CRC64;

Query Match 100.0%; Score 1308; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-107;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPPFGVHVSQALGTLFLFLAASLSAQNCGWDSPTCTEGVSVSGWNTVMSCSISNA 60
Db 1 MOTCLAPPFGVHVSQALGTLFLFLAASLSAQNCGWDSPTCTEGVSVSGWNTVMSCSISNA 60
Qy 61 FSHVNIKLRAHQESAI FNEVAPGFYFSDRGWQLOVQGGVAQLVTKGARDSHAGLYMHV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGFYFSDRGWQLOVQGGVAQLVTKGARDSHAGLYMHV 120
Qy 121 GHORNNRQVTLVSGAEPSQADPTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPSQADPTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Qy 181 KFFLLEPMKVAAALRAGAQOGLSRASAEI LTPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Db 181 KFFLLEPMKVAAALRAGAQOGLSRASAEI LTPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 2

Q921W8 Q921W8 PRELIMINARY; PRT; 192 AA.
ID Q921W8
AC Q921W8
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CDNA sequence BC010462.
GN Name=BC010462;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N. TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N. TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010462; AAH10462.1; -;
DR MGD; MGI:2384805; BC010462.
DR SMART; SM00409; IG; 1.
SQ SEQUENCE 192 AA; 21438 MW; 13D8989D3F07AE1F CRC64;

Query Match 29.2%; Score 382; DB 2; Length 192;
Best Local Similarity 43.6%; Pred. No. 9.3e-26;
Matches 85; Conservative 31; Mismatches 55; Indels 24; Gaps 6;
Qy 1 MOTCLAPPFGVHVSQALGTLFLFLAASLSAQNCGWDSPTCTEGVSVSGWNTVMSCSISNA 60
Db 1 MMTCP-SVPAIPTLWLFSLILLVLSNAQNKSHDNFICTEGILSVPRGNPAVMTCHSNT 59
Qy 61 FSHVNIKLRAHQESAI FNEVAPGFYFSDRGWQLOVQGGVAQLVTKGARDSHAGLYMHV 120
Db 60 FTDVITQLSANGKDKTIFDKKPGQNFSGRWELQVQGGQALVTKDQDDHTGIYWLQHL 119
Qy 121 GHORNNRQVTLVSGAEPSQ--APTGFV-----PVP-----AVVTAVFILL 160
Db 120 GRORCYKNTLNT--LEPSNEDKVPDPTLTFTSPFDHAKSPIEGKPGTLGVITVIFLG 177
Qy 161 VA-LVMFWAYRCRCS 174
Db 178 VAGFITFIYRHRSS 192

RESULT 3

Q6P781 Q6P781 PRELIMINARY; PRT; 218 AA.
ID Q6P781
AC Q6P781
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC72571.
GN Name=MGC72571;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

[illegible]

RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN=CZECH II.; TISSUE=Mammary tumor;
 RC Strauberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC020159; AAH20159.1; -;
 DR MGD; MGI:1929083; Sectm1.
 DR SMART; SM00409; IG; 1.
 SQ SEQUENCE 211 AA; 23433 MW; 9B6FB436592BB2B8 CRC64;

 Query Match 25.0%; Score 327.5; DB 2; Length 211;
 Best Local Similarity 40.0%; Pred. No. 6.9e-21;
 Matches 78; Conservative 24; Mismatches 70; Indels 23; Gaps 3;

 QY 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNAPSHVNIKL 69
 DB 10 GLFPRMLWALLLLAASLNAYNHVMDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
 QY 70 AHGQESAINEVAPGYFSRDGWLQVOGVGAQLVKGARDSHAGLYMWHLVGHORNQV 129
 DB 70 TSEKTSIIFNQTPGNTSKOSQWHLIQQGQALVITDAQKHSGEYWKURGLQAEFKNF 129
 QY 130 TLEVSGAEPQAPDTGFWPVP-----AVVTAVFILLVA-----LVMFAWYRCRCQQ 177
 DB 130 NLIVNAADRKTELDLPVTKVDPKSTAVTVEVIIATTTIITGIGVFWY-----181
 QY 178 REKKFFLLEPQMKVA 192
 DB 182 ---KQFPVAPQIQMS 193

 RESULT 6
 SCTM_MOUSE STANDARD; PRT; 212 AA.
 AC Q9JL59;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Secreted and transmembrane protein 1 precursor.
 GN Name=Secretm1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A., AND INTERACTION WITH CD7.
 RC TISSUE=Colon;
 RX MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
 RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
 RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
 RL protein.";
 RL J. Biol. Chem. 275:3431-3437 (2000).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Skalska D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- SUBUNIT: Interacts with CD7.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also
 CC found as secreted (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF210700; AAF30406.1; -;
 DR EMBL; BC010805; AAH10805.1; -;
 DR MGD; MGI:1929083; Sectm1.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR007110; IG-like.
 DR SMART; SM00409; IG; 1.
 KW Signal; Transmembrane.
 FT SIGNAL 1 28 Potential.
 FT CHAIN 29 212 Secreted and transmembrane protein 1.
 FT DOMAIN 29 160 Extracellular (Potential).
 FT TRANSMEM 161 181 Potential.
 FT DOMAIN 182 212 Cytoplasmic (Potential).
 FT DISULFID 38 55 Potential.
 FT CARBOHYD 56 56 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 85 85 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 114 114 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 130 130 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 212 AA; 23477 MW; 75113E877A2C5B87 CRC64;

 Query Match 24.9%; Score 326; DB 1; Length 212;
 Best Local Similarity 37.8%; Pred. No. 9.4e-21;
 Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

 QY 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNAPSHVNIKL 69
 DB 10 GLFPRMLWALLLLAASLNAYNHVMDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
 QY 70 AHGQESAINEVAPGYFSRDGWLQVOGVGAQLVKGARDSHAGLYMWHLVGHORNQV 129
 DB 70 TSEKTSIIFNQTPGNTSKOSQWHLIQQGQALVITDAQKHSGEYWKURGLQAEFKNF 129
 QY 130 TLEVSGAEPQAPDTGFWPVP-----AVVTAVFILLVA-----LVMFAWYRCRCQQ 176
 DB 130 NLIVNAADRKTELDLPVTKVDPKSTAVTVEVIIATTTIITGIGVFWY-----182
 QY 177 REKKFFLLEPQMKVA 192
 DB 183 ---KQFPVAPQIQMS 194

 RESULT 7
 Q9D966 PRELIMINARY; PRT; 212 AA.
 ID Q9D966
 AC Q9D966;
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Mus musculus (1 day old male pancreas cDNA, RIKEN full-length enriched
 DE library, clone:1810003C24 product:secreted and transmembrane 1, full
 DE insert sequence).
 DE Name=Secretm1;
 GN Name=Secretm1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Tanaka S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki T., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK029082; BAC26284.1; -.
DR MGD; MGI:1929083; Sectm1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
KW Transmembrane.
SQ SEQUENCE 212 AA; 23559 MW; C863064DCE36EB7E CRC64;

Query Match 24.8%; Score 324; DB 2; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.4e-20;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

Qy 10 GHVSQALGTLFLAASLAQNEGDSPTICTEGVSVSGENTVMSCNISAFSHVNIKL 69
Db 10 GLFPRMLWALLLAASLNAYNHVNDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
Qy 70 AHGQESAFNEVAGYFSRDQWQLOVQGGVAQLVIKGARDSHAGLYVHWHVGHORNQOV 129
Db 70 TSKKTSIFNKTPEGNYSKDSQWLHIQGGQALVITDAQGHKSGEYWMKLRGFAEPKNF 129
Qy 130 TLEVSAGPEQSPADPTGFVPV-----AVTAVFILLVALVMPFAYWYRCRCSQ 176
Db 130 NLIWNAORQKTEDLPVTKVDPKPTAVRTEVIIIAITIIITGIVGFVWY----- 182
Qy 177 RREKKFFLEPQMKVA 192
Db 183 ----KQPPVAPQIQMS 194

RESULT 9
ID Q52289 PRELIMINARY; PRT; 550 AA.
AC Q52289;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE UxPA protein.
GN Name=uxpa;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WCS358;
RX MEDLINE=9618681; PubMed=8602167; DOI=10.1007/s004380050102;
RA de Groot A., Krizger J.J., Filloux A., Tomassen J.;
RT "Characterization of type II protein secretion (xcp) genes in the
RT plant growth-stimulating Pseudomonas putida, strain WCS358.";
RL Mol. Gen. Genet. 250:491-504(1996).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=WCS358;
RA de Groot A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; X81085; CAA56977.1; -.
DR PIR; S64725; S64725.
SQ SEQUENCE 550 AA; 59118 MW; 684D79D0D279868B CRC64;

Query Match 7.9%; Score 103.5; DB 2; Length 550;
Best Local Similarity 23.2%; Pred. No. 1.4;
Matches 72; Conservative 42; Mismatches 118; Indels 79; Gaps 16;

Qy 2 QTCLAPFGHVSOA-LGTLFLAA-----SLSAQN-EGWDSPTICTEGVSVS 46
Db 90 QLAPLLDAGRADANLGGYAVLGLLEQLGEGAGAGNSLTLENGQGWNG-----SGLAYLT 145
Qy 47 WGMTVMSCNISAFSHVNIKLRA--HGQESAFNEVAGYFSR-----DWGQ 92
Db 146 QGEGVQSGQLLSEARVSDERVLWPQSAALYRQASATTLGAGLADEQRQALGLEPLQ 205
Qy 93 LQVQGGVAQLVIKG-----ARDSHAGLYVWH-----LVGHORNQOVTLVSGAPEQSAP-- 142
Db 206 LFERGG-ARIADVGVTDPAQADQKASLKQWYQALLPVFQOARREADLVVALADVGTGPGL 264
Qy 143 -----DTGFVPVAVVTAVFILLVALVMPF-----YRCRC-----SQ 176
Db 265 WLAERLPAIDLLLCARGQDLWPTFVQATQASGRVP-VLPAGCRGSGAFRLRCQOVAGOW 323
Qy 177 RREKKFF-----LLEPQMKVAALRAGAQQGLSASAEELWTPDSEPTPR-PLALVFKPSPL 230
Db 324 QFSGRFFPTFEQTLSPAAQVRAQLOALNQQRAGHAAML--DOPLARAPQALWRRDTRG 381
Qy 231 GALELLSPQPL 241
Db 382 GSWDRLLHQAL 392

RESULT 10
ID SHS1_BOVIN STANDARD; PRT; 506 AA.
AC O46631; O46632;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
DE (SHP substrate-1) (SHPs-1) (inhibitory receptor SHPS-1) (Signal-
GN Name=PTNS1; Synonyms=MYD1, SHPS1, SIRP;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS SBR-23; ALA-28; LEU-61; ARG-70;
RP HIS-120; 125-GLN; 127-GLY; 129-HIS; 132-VAL; ASN-145; VAL-153;
RP ASP-203; ARG-261; LEU-302; LEU-316; ARG-337; ASN-367; LEU-422; PHE-429
RP AND GLU-433.
RC STRAIN=Friesian; TISSUE=Peripheral blood;
RX MEDLINE=98143722; PubMed=9485180;
RX DOI=10.1002/(SICI)1521-4141(199801)28:01<1::AID-IMMU1>3.3.CO;2-M;
RA Brooke G.P., Parsons K.R., Howard C.J.;
RT "Cloning of two members of the SIRP alpha family of protein tyrosine
RT phosphatase binding proteins in cattle that are expressed on monocytes
RT and a subpopulation of dendritic cells and which mediate binding to
RT CD4 T cells.";
RL Eur. J. Immunol. 28:1-11(1998).
CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
CC as docking protein and induces translocation of PTPN6, PTPN11 and
CC other binding partners from the cytosol to the plasma membrane.
CC Supports adhesion of cerebellar neurons, neurite outgrowth and
CC glial cell attachment. May play a key role in intracellular
CC signaling during synaptogenesis and in synaptic function. Involved
CC in the negative regulation of receptor tyrosine kinase-coupled

```



```
Db 348 ----VCERDDHTSAQLTVSPRVVFKMPSGLSTVVAEBGGEATFCQVVSVDVAVV----- 399
Oy 166 FAWYRCRCSQORREKKFFLEPQMKVAALRAGAQQGLSRASAE-----WTPDSPTPRP 220
Db 400 --WFRDGA-----LLQSEKFAISQSGASHLSITSLVLEDAQOITVEABGASS 447
Oy 221 LALVFKPSPGLGALELLSPQ 239
Db 448 AALRVREAPVLFKKKLEPQ 466

RESULT 14
Q9V5F7 PRELIMINARY; PRT; 894 AA.
AC Q9V5F7;
DT 01-MAY-2000 (TrEMBLrel. 12, Created)
DT 01-OCT-2002 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG2292-PA.
GN ORFNames=CG2292;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.E., Ferraz C., Ferrier S., Fleischmann W.,
RA Fowler C., Gabriellian A.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Stong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
PT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
```

```
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003831; AAF58853.2;
DR FlyBase; FBGN0033479; CG2292.
DR InterPro; IPR007070; PIGN.
DR Pfam; PF04987; PIGN; 1.
SQ SEQUENCE 894 AA; 101828 MW; 9BD68206E9595187 CRC64;

Query Match 7.4%; Score 97; DB 2; Length 894;
Best Local Similarity 22.7%; Pred. No. 8.8;
Matches 62; Conservative 33; Mismatches 88; Indels 90; Gaps 15;

Oy 9 PGHVSQALGTLLFLAASLSAQNEGWS-PICTEGWSVS-----WGENTVMSNCISNAFS 62
Db 95 PGHTLIAG-----LYEDPSAVLGKWKNPIDFDFVFNRSOTYANGANDVL-----NVFS 145
Oy 63 HVN-----IKLRAHQESAI FNEVAPGY--FSRDGWLQVQGGVAQLVIKGARDSHAGLYM 116
Db 146 HVSNGGBINLRFYNHDL-----DFSPGYDAYEEDSEWFKRVKLLQLQKREALQRAQNVVFF 201
Oy 117 WHLVGHORNNRQVTLVEYSG-AEQSAPDTGFWPVPVAVTAVFILLVALVMFVWRCRCQ 175
Db 202 LHLG-----LDTAGHVHKPGAP-----KFRRTL 225
Oy 176 QRREK-----KFFLEPQMKVAALRAGAQQGLSRASAE-----LW-TPDS 214
Db 226 EKTEKGVYVYQIFERFVFPDKRTAYLLIT-ADHGWTDSGAHSGSPHETDTPFMLWGAGAS 284
Oy 215 EPTPRPALVFKPS-----PLGALELLSPQL 241
Db 285 RAVPKPGCGTFMPNNEGPAHLEQAQLTPL 317

RESULT 15
Q88F09 PRELIMINARY; PRT; 550 AA.
AC Q88F09;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005. CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 06:04:49 ; Search time 122 Seconds
(without alignments)
786.201 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCLAPGHVSQALGTLL.....PLGALLELLSPQLFPYAADP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	248	2 AAW48811	Aaw48811 K12 prote
2	1308	100.0	248	2 AAY21846	Aay21846 Human sig
3	1308	100.0	248	4 AAB36658	Aab36658 Human K12
4	1308	100.0	248	8 ADJ75366	Adj75366 Marker ge
5	1308	100.0	248	8 ADP24664	Adp24664 PRO polyp
6	916.5	70.1	183	6 AAO29897	Aao29897 Human org
7	720.5	55.1	149	7 ADB36336	Adb36336 Human imm
8	719.5	55.0	162	6 AAO29896	Aao29896 Human org
9	492	37.6	101	6 AAO29895	Aao29895 Human org
10	326	24.9	212	4 AAB36660	Aab36660 Mouse K12
11	326	24.9	212	8 ADJ76167	Adj76167 Marker ge
12	325	24.8	212	8 ADJ76168	Adj76168 Marker ge
13	294	22.5	107	6 AAO29898	Aao29898 Human org
14	286	21.9	55	3 AAB34671	Aab34671 Gene 35 h
15	279	21.3	55	3 AAB34672	Aab34672 Human sec
16	163	12.5	30	3 AAB34673	Aab34673 Gene 35 h
17	156	11.9	43	7 ADB36337	Adb36337 Human imm
18	100	7.6	874	4 ABB59013	Abb59013 Drocephil
19	96	7.3	435	6 ABR42173	Abr41273 Human DIT
20	96	7.3	442	4 AAB95569	Aab95569 Human pro
21	95.5	7.3	506	5 ADI16910	Adi16910 Cow NOVX
22	91	7.0	6620	7 ADJ70485	Adj70485 Human hea
23	91	7.0	7968	6 ADJ70487	Adj70487 Human ser
24	90.5	6.9	1171	7 ABO63781	Abo63781 Klebsell
25	89	6.8	435	4 AAB61190	Aab61190 Mature hu

Abo32633 Secreted
Adb90720 Human INT
Adf71455 Human INT
Aag10281 Human pol
Aab61188 Human INT
Abo32631 Secreted
Adb90718 Human INT
Adf71453 Human INT
Adq10279 Human pol
Adn99576 Novel hum
Abul5711 Protein e
Aar12427 Hybrid Pc
Abo79025 Pseudomon
Adq59388 Human can
Adq76631 Human but
Adq76633 Human but
Aabm83615 Human dia
Adq67749 Novel hum
Abr58656 Human can
Abp69627 Human pol

ALIGNMENTS

RESULT 1
AAW48811
ID AAW48811 standard; protein; 248 AA.
XX AC AAW48811;
XX DT 26-OCT-1998 (first entry)
XX DE K12 protein.
XX KW DNA probe; CD7 Hs1 DNase hypersensitive site; mRNA northern blot;
XX KW human erythroleukemic; HEL; K562 cell line; Clone; breast cancer;
XX KW ovarian cancer; malignant; tumour.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 239 /note= "encoded by CCAA"
XX FT
XX PN WO9822502-A1.
XX PD 28-MAY-1998.
XX PF 24-NOV-1997; 97WO-US021517.
XX PR 22-NOV-1996; 96US-00755559.
XX PA (UYDU-) UNIV DUKE.
XX PI Kaufman RE, Slentz-Keeler KA;
XX WPI; 1998-312415/27.
XX DR N-PSDB; AAV32446.
XX PS New isolated K12 protein gene - which is over expressed in certain
XX PS neoplastic cells, used to develop products for tumour detection and
XX PS treatment.
XX PS Claim 3; Fig 1; 44pp; English.
XX CC This present sequence represents the K12 protein, the gene for which has
XX CC been located on chromosome 17q25. To obtain this genes cDNA sequence a
XX CC 500 bp DNA probe, which can be located just upstream of the CD7 Hs1 DNase
XX CC hypersensitive site, was used against a mRNA northern blot. From this a
XX CC 1.8kb transcript was detected in the human erythroleukemic cell line HEL.
XX CC The probe was then used to screen a human erythroleukemic cell line K562
XX CC cDNA library, from which several clones were identified and isolated that

CC constituted a 1.8kb cDNA. This cDNA was designated K12 and was found to
 CC have a single open reading frame as well as being in the same orientation
 CC as CD7. The K12 gene was found to be expressed in both breast and ovarian
 CC cancer cells at a much higher level than any other malignant or normal
 CC tissue that was examined, thus enabling the K12 to be a useful protein in
 CC tumour detection and treatment

SQ Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.8e-133;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSNCISNA 60
 DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHGOESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMWHLV 120
 DB 61 FSHVNIKLRAHGOESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMWHLV 120

QY 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2

AA121846
 ID AA121846 standard; protein; 248 AA.

AC AA121846;

XX 20-SEP-1999 (first entry)

XX Human signal peptide-containing protein (SIGP) (clone ID 1747327).

XX Signal-peptide containing protein; SIGP; human; cancer; immune response;
 KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
 KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
 KW asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease;
 KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
 KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
 KW Grave's Disease; hyperesinophilia; irritable bowel syndrome; infection;
 KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
 KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
 KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.

XX Homo sapiens.

XX WO9333981-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US027598.

XX 31-DEC-1997; 97US-00002485.

XX (INCY-) INCYTE PHARM INC.

XX Lal P, Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
 PI Shah P;

XX WPI; 1999-430242/36.

XX N-PSDB; AA82081.

PT Human signal-peptide containing protein coding sequences used to treat
 XX cancer and immune responses.

PS Claim 1; Page 79-80; 99pp; English.

XX The invention provides human signal-peptide containing proteins (SIGP)
 CC (AA121841-855) and polynucleotides (AA121841-855) encoding the proteins. A
 CC host cell containing a vector comprising SIGP DNA can be used to produce
 CC the SIGP protein. The SIGP protein can be used, in conjunction with a
 CC pharmaceutical carrier to treat or prevent a cancer. An antagonist of the
 CC SIGP protein can be used to treat or prevent a cancer or an immune
 CC response. The cancers that can be treated or prevented include sarcomas,
 CC adenocarcinomas, leukemia's, lymphomas, melanomas, teratocarcinomas,
 CC myelomas and cancers of the adrenal gland, bladder, bone, bone marrow,
 CC brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract,
 CC heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis,
 CC prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and
 CC uterus. The immune responses that can be treated or prevented include
 CC AIDS, Addison's disease, adult respiratory distress syndrome, allergies,
 CC anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's
 CC disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes
 CC mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's
 CC disease, gout, hyperesinophilia, irritable bowel syndrome, lupus
 CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or
 CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
 CC polyomyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and
 CC autoimmune thyroiditis, complications of cancer, infections, and trauma

SQ Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 2; Length 248;

Best Local Similarity 100.0%; Pred. No. 1.8e-133;

Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSNCISNA 60

DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHGOESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMWHLV 120

DB 61 FSHVNIKLRAHGOESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMWHLV 120

QY 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180

DB 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLELLSPQ 240

DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248

DB 241 LFPYAADP 248

RESULT 3

AAB36658

ID AAB36658 standard; protein; 248 AA.

XX AAB36658;

XX 13-MAR-2001 (first entry)

XX Human K12 protein sequence SEQ ID NO:4.

XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
 KW identification; inhibiting T cell proliferation; HIV; infection;
 KW activating natural killer cell proliferation; leukaemia; lymphoma;
 KW sepsis; graft versus host disease; autoimmune disease; arthritis;
 KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
 KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
 KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
 KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;

KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis.
OS Homo sapiens.
XX WO20007333-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014612.
XX 28-MAY-1999; 99US-0136450P.
XX (IMMV) IMMUNEX CORP.
XX Lyman SD, Fanslow WC;
XX WPI; 2001-061511/07.
DR N-PSDB; AAC88152.
XX
XX Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.
XX
XX Claim 2; Page 38-39; 42pp; English.
XX
XX The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatitis, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC human K12 protein, which is given in the exemplification of the present
CC invention
XX
SQ Sequence 248 AA;
Query Match 100.08; Score 1308; DB 4; Length 248;
Best Local Similarity 100.08; Pred. No. 1.8e-133;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MOTCLAPFGHVHVSQALGTLFLFLASLSAQNGWDSPICTEGVSVSGWNTVMSGNTSNA 60
DB 1 MOTCLAPFGHVHVSQALGTLFLFLASLSAQNGWDSPICTEGVSVSGWNTVMSGNTSNA 60
OY 61 FSHVNIKLRAHQESAIENEVAPGYFSDRGVQVGGVQVQVQVQVQVQVQVQVQVQV 120
DB 61 FSHVNIKLRAHQESAIENEVAPGYFSDRGVQVGGVQVQVQVQVQVQVQVQVQVQV 120
OY 121 QHQRNRQVTLVSCAEQSPADPTGFWPVAWTAFTVILLVAFVWYRCRCQQRREK 180
DB 121 QHQRNRQVTLVSCAEQSPADPTGFWPVAWTAFTVILLVAFVWYRCRCQQRREK 180
OY 181 KFFLELPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPLCALSLSPQ 240
DB 181 KFFLELPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPLCALSLSPQ 240

OY 241 LFPYAADP 248
DB 241 LFPYAADP 248
RESULT 4
ID ADJ75366
XX ADJ75366 standard; protein; 248 AA.
XX AC ADJ75366;
XX 20-MAY-2004 (first entry)
XX Marker gene related amino acid sequence SEQ ID NO:618.
XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX Homo sapiens.
XX EPI394274-A2.
XX 03-MAR-2004.
XX 04-AUG-2003; 2003EP-00254857.
XX 06-AUG-2002; 2002JP-00229312.
XX 20-MAR-2003; 2003JP-00077212.
XX (GENO-) GENOX RES INC.
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
XX WPI; 2004-193155/19.
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
XX Example 11; SEQ ID NO 618; 241pp; English.
XX
XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX

XX SQ Sequence 149 AA;

Query Match 55.1%; Score 720.5; DB 7; Length 149;
 Best Local Similarity 97.2%; Pred. No. 8e-70; 2; Indels 1; Gaps 1;
 Matches 138; Conservative 1; Mismatches 2;

Qy 1 MOTCPALPFGHVSQLGTLFLAASLSAQNEGWDSPICTEGVSVSMGENTVMSNCISNA 60
 Db 1 MOTCPALPFGHVSQLGTLFLAASLSAQNEGWDSPICTEGVSVSMGENTVMSNCISNA 60

Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHLV 120
 Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHLV 120

Qy 121 GHQNNRQVTLVSGAE-POSA 141
 Db 121 GHQNNRQVTLVSGAAVPSNA 142

RESULT 8
 AAO29896
 ID AAO29896 standard; protein; 162 AA.
 AC AAO29896;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA)-9.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 DR WPI; 2003-457603/43.
 DR N-PSDB; AAL60549.
 XX
 PT New organelle-associated proteins and polymucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 PS Claim 1; Page 166; 194pp; English.
 PS
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful

CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX
 SQ Sequence 162 AA;

Query Match 55.0%; Score 719.5; DB 6; Length 162;
 Best Local Similarity 89.1%; Pred. No. 1.2e-69;
 Matches 139; Conservative 0; Mismatches 6; Indels 11; Gaps 1;

Qy 1 MOTCPALPFGHVSQLGTLFLAASLSAQNEGWDSPICTEGVSVSMGENTVMSNCISNA 60
 Db 1 MOTCPALPFGHVSQLGTLFLAASLSAQNEGWDSPICTEGVSVSMGENTVMSNCISNA 60

Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHLV 120
 Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHLV 120

Qy 121 GHQNNRQVTLVSGAE-POSA 145
 Db 121 GHQNNRQVTLVSGAAVPSNA 156

RESULT 9
 AAO29895
 ID AAO29895 standard; protein; 101 AA.
 AC AAO29895;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA)-8.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;

XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60548.
 XX
 XX New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 XX Claim 1; Page 165-166; 194pp; English.
 XX
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatic), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX
 XX Sequence 101 AA;
 SQ
 Query Match 37.6%; Score 492; DB 6; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.5e-45;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MOTCLAPPGHVQSALGTLFLAASLSAQNGWDSPICTEGVSVSGWNTVMSCNTSNA 60
 Db 1 MOTCLAPPGHVQSALGTLFLAASLSAQNGWDSPICTEGVSVSGWNTVMSCNTSNA 60
 Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQ 92
 Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQ 92
 RESULT 10
 AAB36660
 ID AAB36660 standard; protein; 212 AA.
 XX
 AC AAB36660;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Mouse K12 protein sequence SEQ ID NO:8.
 XX
 KW Mouse; CD7; K12; cognate ligand; cluster of differentiation; cancer;
 KW identification; inhibiting T cell proliferation; HIV; infection;
 KW activating natural killer cell proliferation; leukaemia; lymphoma;
 KW sepsis; graft versus host disease; autoimmune disease; arthritis;
 KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
 KW scleroderma; psoriasis; atopic dermatitis; type 1 diabetes mellitus;
 KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
 KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
 KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
 KW atypical dermatitis; contact dermatitis; eczematous dermatitis;
 KW seborrheic dermatitis; rhinitis.
 XX
 OS Mus sp.
 XX
 PN WO200073333-A2.
 XX
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014612.
 PF
 XX 28-MAY-1999; 99US-0136450P.
 PR
 XX (IMMV) IMMUNEX CORP.
 FA

XX Lyman SD, Fanslow WC;
 XX WPI; 2001-061511/07.
 DR N-PSDB; AAC88154.
 XX
 XX Stimulating intracellular signaling of CD7 comprises contacting a cell
 PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
 PT to inhibit T cell proliferation and/or activate natural killer cell
 PT proliferation.
 XX
 XX Claim 32; Page 42; 42pp; English.
 XX
 XX The present invention describes a method for stimulating (S) the
 CC intracellular signalling of CD (cluster of differentiation) 7 comprising
 CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
 CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
 CC proliferation and/or activating NK (natural killer) cell proliferation
 CC and/or inducing NK toxicity in a mammal which involves administration of
 CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
 CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
 CC bacterial and viral infections, mediated by CD7. In the case of treating
 CC T cell leukaemia the soluble K12 protein is covalently attached to a
 CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
 CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
 CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
 CC psoriasis, atopic dermatitis, type 1 diabetes mellitus, Hashimoto's
 CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
 CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
 CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical
 CC dermatitis, contact dermatitis, other eczematous dermatitis, seborrheic
 CC dermatitis, and rhinitis is also treated by administering a K12
 CC antagonist (neutralising antibody). The present sequence represents the
 CC mouse K12 protein, which is given in the exemplification of the present
 CC invention
 XX
 XX Sequence 212 AA;
 SQ
 Query Match 24.9%; Score 326; DB 4; Length 212;
 Best Local Similarity 37.8%; Pred. No. 1.2e-26;
 Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
 Qy 10 GHVSQALGTLFLAASLSAQNGWDSPICTEGVSVSGWNTVMSCNTSNAFHVNIKL 69
 Db 10 GLFPRMLWALLLLAASLSAHNDVWDEPCCTEHEVSVNRGSRVNVACISNNLRDVTIELV 69
 Qy 70 AHGOESAIENEVAPGYFSRDGQQLVQGGVAQLVKGARDSHAGLYMMHLYGHQNRNQV 129
 Db 70 TSEKTSIIFNHTPPGNYKSDSWQLHIQGVQALVITDAQDKHSGNYSKWLHGFQAEKFN 129
 Qy 130 TLEVSAGBPOSAPDTGFWVP-----AVTAVFILLVALVMFAWYRCRSQQ 176
 Db 130 NLTVNAADROKTDLPVTKVPDRPPTAVRTEVIIIAITTIITIGIVFWY----- 182
 Qy 177 RREKKFPLEPOMKVA 192
 Db 183 -----KOPPVAPQIQMS 194
 RESULT 11
 ADJ76167
 ID ADJ76167 standard; protein; 212 AA.
 XX
 AC ADJ76167;
 XX
 XX 20-MAY-2004 (first entry)
 DT
 XX Marker gene related amino acid sequence SEQ ID NO:1419.
 DE
 XX bronchial asthma; chronic obstructive pulmonary disease;
 XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker.
 KW
 XX


```

OS Mus musculus.
XX EP1394274-A2.
XX
XX PD 03-MAR-2004.
XX
XX PF 04-AUG-2003; 2003EP-00254857.
XX
XX PR 06-AUG-2002; 2002JP-00229312.
XX
XX PR 20-MAR-2003; 2003JP-00077212.
XX
XX PA (GENO-) GENOX RES INC.
XX
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX
XX DR WPI; 2004-193155/19.
XX
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a
XX healthy subject.
XX
XX PS Claim 16; SEQ ID NO 1419; 241pp; English.
XX
XX CC The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.
XX
XX SQ Sequence 212 AA;

Query Match          24.9%; Score 326; DB 8; Length 212;
Best Local Similarity 37.8%; Pred. NO. 1.2e-26;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

QY 10 GHVSQALGTLFLAASLAQNEGWDSPICTGCVSVSGENTVMSCHNISAFSHVNIKL 69
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 GLFFRMLWALLLLAASLNANHDVMDPECCCTHEYSVNRGSRVNMACHNNLRDVTIEL 69
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 70 AHGESATFNEVPYFGRDCWQLOVQGVQALVTKGARDSDHAGLYMWHLVGHQRNNQV 129
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 70 TSEKTSIFNHTPPGNYSKDSQWLIHQVQALVITDAQDKHSGNSKWLHGFQAEKFN 129
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 130 TLEVSAGBPQAPDPTGFVFP-----AVVTAVPILVALVMPFAWYRCRSQQ 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 130 NLTVNAADRKQTDLPVTKPKDPKPTAVRTEVITIIIAITIIITIGVFWY----- 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 177 RREKKFFLLEPQMKVA 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 183 ----KQFPVAPQIQMS 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
ADJ76168
ID ADJ76168 standard; protein; 212 AA.
XX
XX AC ADJ76168;
XX
XX DT 20-MAY-2004 (first entry)
XX
XX DE Marker gene related amino acid sequence SEQ ID NO:1420.
XX
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX gene therapy; marker.
XX
XX OS Mus musculus.
XX
XX PN EP1394274-A2.
XX
XX PD 03-MAR-2004.
XX
XX PF 04-AUG-2003; 2003EP-00254857.
XX
XX PR 06-AUG-2002; 2002JP-00229312.
XX
XX PR 20-MAR-2003; 2003JP-00077212.
XX
XX PA (GENO-) GENOX RES INC.
XX
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX
XX DR WPI; 2004-193155/19.
XX
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a
XX healthy subject.
XX
XX PS Claim 16; SEQ ID NO 1420; 241pp; English.
XX
XX CC The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.
XX
XX SQ Sequence 212 AA;

```



```
Query Match      24.8%; Score 325; DB 8; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.5e-26;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

QY 10 GHVSQALGTLFLAASLQAQNEGWDSPICTEGVSVSGENTVMSCNISNAFHVNIKL 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 10 GLFFRMLWALLLLAASLNAYNHVDKPCCTEHEVSVNRGSRVYVMACTNSNLLRDVTIELV 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 70 AHGQSAIFNEVAPGFRDQWQLOVQGGVAQLVIKGAARDSHAGLYMHLVGHQRNNROV 129
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 70 TSKKTSIFNQTPPGNYSKDSQLHIQGGQALVITDAQKHSYWKLRGFOAEFKNF 129
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 130 TLEVSGAEPQSDAPDTGFVPVP-----AVTAVFILLVALVNFAMWYRCRCRCSQQ 176
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 130 NLIVNAADRQKTEDELVTVKVPDKPTAVRTAVTEVIIIIATIIITIGIGVFVWY----- 182

QY 177 RREKKFFLLEPOMKVA 192
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 183 ----KQFPVAPQIQMS 194

RESULT 13
AAO29898
ID AAO29898 standard; protein; 107 AA.
XX AC AAO29898;
XX DT 03-SEP-2003 (first entry)
XX DE Human organelle-associated protein (ORGA)-11.
XX KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
XX KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
XX KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
XX KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
XX KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
XX KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
XX KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
XX KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
XX KW infection; urethritis; prostatitis.
XX OS Homo sapiens.
XX PN WO2003044171-A2.
XX PD 30-MAY-2003.
XX PF 15-NOV-2002; 2002WO-US036807.
XX PR 16-NOV-2001; 2001US-0332384P.
XX PR 13-DEC-2001; 2001US-0341187P.
XX PR 23-JAN-2002; 2002US-0351151P.
XX PR 27-FEB-2002; 2002US-0360269P.
XX PR 05-APR-2002; 2002US-0370637P.
XX PR 14-JUN-2002; 2002US-0388946P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
XX PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
XX PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
XX PI Hafalia AJA, Bulloch S;
XX DR WPI; 2003-457603/43.
XX DR N-PSDB; AAL60551.
XX PT New organelle-associated proteins and polynucleotides, useful for
XX PT diagnosing, treating and/or preventing cell proliferative, reproductive,
XX PT gastrointestinal, neurological, urologic, and renal disorders.
XX PS Claim 1; Page 167; 194pp; English.
XX
```

```
CC The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide esophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein
XX
SQ Sequence 107 AA;

Query Match      22.5%; Score 294; DB 6; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPLAFPGHVSOALGTLFLAASLQAQNEGWDSPICTEGVSVSGENTVMS 55
   |||||
DB 1 MOTCPLAFPGHVSOALGTLFLAASLQAQNEGWDSPICTEGVSVSGENTVMS 55
   |||||

RESULT 14
AAB34671
ID AAB34671 standard; protein; 55 AA.
XX AC AAB34671;
XX DT 26-JAN-2001 (first entry)
XX DE Gene 35 human secreted protein homologous amino acid sequence #155.
XX KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
XX KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
XX KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
XX KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
XX KW cerebrovascular disorder; angiogenesis; nervous system disorder;
XX KW ocular disorder; wound healing; skin aging; food additive; preservative.
XX OS Homo sapiens.
XX PN WO2000056751-A1.
XX PD 28-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US006013.
XX PR 19-MAR-1999; 99US-0125360P.
XX PR 11-JUN-1999; 99US-0138626P.
XX PR 03-DEC-1999; 99US-0168662P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX DR WPI; 2000-579482/54.
XX PT Isolated nucleic acid molecule encoding a human secreted protein is used
XX PT in preventing, treating or ameliorating a medical condition.
XX PS Disclosure; Page 412; 419pp; English.
XX CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
XX CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
XX CC AAB34686 represent human secreted polypeptide sequences and proteins
XX CC homologous to them, which are given in the exemplification of the present
XX CC invention. Human secreted proteins have activities based on the tissues
XX CC and cells the genes are expressed in. Example of activities include:
```


CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytosatic; cardiant; vasotropic; cerebroprotective; nontropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention

XX SQ Sequence 55 AA;

Query Match 21.9%; Score 286; DB 3; Length 55;

Best Local Similarity 100.0%; Pred. No. 3.9e-23;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 LRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLSPQLFPFYAADP 248

Db 1 LRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLSPQLFPFYAADP 55

RESULT 15

AAB34672

ID AAB34672 standard; protein; 55 AA.

XX AC AAB34672;

XX DT 26-JAN-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 35 SEQ ID NO:156.

XX KW Human; secreted protein; diagnosis; immunosuppressive; antirheumatic;
 KW antirheumatic; antiproliferative; cytosatic; cardiant; vasotropic;
 KW cerebroprotective; nontropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative.

XX OS Homo sapiens.

XX OS WO2000056751-A1.

XX PN 28-SEP-2000.

XX PD 09-MAR-2000; 2000WO-US006013.

XX PR 19-MAR-1999; 99US-0125360P.

XX PR 11-JUN-1999; 99US-0138626P.

XX PR 03-DEC-1999; 99US-0168662P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Rosen CA, Ruben SM, Komatsoulis G;

XX XX WPI; 2000-579482/54.

XX DR Isolated nucleic acid molecule encoding a human secreted protein is used
 XX PT in preventing, treating or ameliorating a medical condition.

XX XX Disclosure; Page 412-413; 419pp; English.

XX PS The polynucleotide sequences given in AAC59738 to AAC59787 encode the

CC

CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; cerebroprotective; nontropic;
 CC cytosatic; cardiant; vasotropic; cerebroprotective; antiproliferative;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention

XX SQ Sequence 55 AA;

Query Match 21.3%; Score 279; DB 3; Length 55;

Best Local Similarity 98.2%; Pred. No. 2.3e-22;

Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 194 LRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLSPQLFPFYAADP 248

Db 1 LRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPLGALLSPQLFPFYAADP 55

Search completed: October 17, 2005, 07:12:16

Job time : 127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 07:12:25 ; Search time 2066 Seconds
(without alignments)

50.036 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MOTCLAPFGHVSQALGTLT.....PLGALELLSPQLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1860064 seqs, 416830855 residues

Total number of hits satisfying chosen parameters: 1860064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	9	US-09-799-777-32 Sequence 32, Appl
2	1308	100.0	248	9	US-09-997-165-4 Sequence 4, Appli
3	1308	100.0	248	14	US-10-080-522-1 Sequence 1, Appli
4	1308	100.0	248	18	US-10-631-467-618 Sequence 618, App
5	326	24.9	212	9	US-09-997-165-8 Sequence 8, Appli
6	326	24.9	212	18	US-10-631-467-1419 Sequence 1419, Ap
7	325	24.8	212	18	US-10-631-467-1420 Sequence 1420, Ap
8	100	7.6	874	20	US-11-097-143-3831 Sequence 3831, Ap
9	98	7.5	969	15	US-10-052-648A-35 Sequence 35, Appl
10	96	7.3	969	15	US-10-052-648A-34 Sequence 34, Appl
11	95.5	7.3	506	15	US-10-072-012-446 Sequence 446, App

12	91.5	7.0	541	16	US-10-719-993-534	Sequence 534, App
13	91.5	7.0	597	15	US-10-425-114-62807	Sequence 62807, A
14	91.5	7.0	597	16	US-10-425-115-279214	Sequence 279214,
15	91	7.0	6620	15	US-10-080-334-290	Sequence 290, App
16	91	7.0	6620	16	US-10-408-765A-2291	Sequence 2291, Ap
17	91	7.0	7968	13	US-10-077-130-5	Sequence 5, Appli
18	89	6.8	435	10	US-09-759-130B-275	Sequence 275, App
19	89	6.8	435	14	US-10-189-123-5	Sequence 5, Appli
20	89	6.8	435	14	US-10-188-495-5	Sequence 5, Appli
21	89	6.8	435	16	US-10-741-790-275	Sequence 275, App
22	89	6.8	455	10	US-09-759-130B-273	Sequence 273, App
23	89	6.8	455	14	US-10-189-123-3	Sequence 3, Appli
24	89	6.8	455	14	US-10-188-495-3	Sequence 273, App
25	89	6.8	455	16	US-10-741-790-273	Sequence 14023, A
26	84.5	6.5	271	14	US-10-156-761-14023	Sequence 43635, A
27	84.5	6.5	289	15	US-10-282-122A-43635	Sequence 24, Appli
28	84.5	6.5	455	16	US-10-322-696-24	Sequence 2, Appli
29	84.5	6.5	455	16	US-10-742-682-2	Sequence 4, Appli
30	84.5	6.5	482	16	US-10-742-682-4	Sequence 8, Appli
31	84.5	6.5	567	18	US-10-737-318-8	Sequence 10, Appl
32	84.5	6.5	792	18	US-10-737-318-10	Sequence 12, Appl
33	84.5	6.5	829	18	US-10-737-318-12	Sequence 6, Appli
34	84.5	6.5	882	18	US-10-737-318-6	Sequence 46, Appl
35	83.5	6.4	199	14	US-10-095-131A-46	Sequence 48, Appl
36	83.5	6.4	199	14	US-10-095-131A-48	Sequence 153, App
37	83.5	6.4	366	14	US-10-091-438-153	Sequence 19, Appl
38	83.5	6.4	368	9	US-09-764-853-678	Sequence 13729, A
39	83.5	6.4	468	17	US-10-626-832-19	Sequence 311858,
40	83.5	6.4	699	15	US-10-369-493-13729	Sequence 107, App
41	83	6.3	340	16	US-10-425-115-311858	Sequence 12, Appl
42	82.5	6.3	286	16	US-10-367-094-107	Sequence 18, Appl
43	82.5	6.3	288	13	US-10-068-215-12	
44	82.5	6.3	288	14	US-10-115-615-12	
45	82.5	6.3	288	16	US-10-748-112-18	

ALIGNMENTS

RESULT 1

US-09-799-777-32

; Sequence 32, Application US/09799777

; Patent No. US20020091244A1

; GENERAL INFORMATION:

APPLICANT: Lal, Preeti

Hillman, Jennifer L.

Corley, Neil C.

Guegler, Karl J.

Baugh, Mariah

Sather, Susan

Shah, Purvi

TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

NUMBER OF SEQUENCES: 154

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/799,777

FILING DATE: 06-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/002,485

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:


```
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-09-799-777-32

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

Qy 61 FSHVNIKLRAHGOESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHWLV 120
Db 61 FSHVNIKLRAHGOESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHWLV 120

Qy 121 GHORNNROVTLVSGAEPOSADPTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Db 121 GHORNNROVTLVSGAEPOSADPTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180

Qy 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
Db 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 2
US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PPT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

Qy 61 FSHVNIKLRAHGOESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHWLV 120
Db 61 FSHVNIKLRAHGOESAIENEVAPGYFSRDGHQLOVQGGVAQLVIKGARDSHAGLYMHWLV 120

Qy 121 GHORNNROVTLVSGAEPOSADPTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Db 121 GHORNNROVTLVSGAEPOSADPTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180

Qy 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
Db 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 3
US-10-080-522-1
; Sequence 1, Application US/10080522
; Publication No. US20030096326A1
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/080,522
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/539,774
; FILING DATE: 31-MAR-2000
; APPLICATION NUMBER: US 09/210,474
; FILING DATE: 14-DEC-1998
; APPLICATION NUMBER: US 08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-645
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-080-522-1

Query Match 100.0%; Score 1308; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 1 MQTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
```


Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Db 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4

US-10-631-467-618
; Sequence 618, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genex Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631.467
; PRIOR FILING DATE: 2003-07-31
; PRIOR FILING DATE: 2003-07-31
; PRIOR FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 618
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-631-467-618

Query Match 100.0%; Score 1308; DB 18; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCPAPPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 1 MOTCPAPPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Db 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 5

US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fenslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF

; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997.165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-997-165-8

Query Match 24.9%; Score 326; DB 9; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.4e-24;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
Qy 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNAFHVNIKL 69
Db 10 GLFPRMLWALLLAASLSNAHNDVMDPCCTEHEVSVNRGSRVVMACNISNNLRDVTIELV 69
Qy 70 AHQGESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLVGHORNNRQV 129
Db 70 TSEKTSIIFNHTPPGNYKSDSWOLHIQGVQAOQLVITDAQDKHSGNYSWKLHGFQAEKFN 129
Qy 130 TLEVSGAEPOSAPDTGFWPVP-----AVVTAVFILLVALVFWAYRCRCSSQ 176
Db 130 NLTVNAADROKTEDLPVTKVPDKPPTAVRTEVIIIIAIIATIIITIGIGVFWY----- 182
Qy 177 RREKKFFLLEPQMKVA 192
Db 183 ----KQPPVAPQIQMS 194

RESULT 6

US-10-631-467-1419
; Sequence 1419, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genex Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631.467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1419
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-631-467-1419

Query Match 24.9%; Score 326; DB 18; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.4e-24;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
Qy 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNAFHVNIKL 69
Db 10 GLFPRMLWALLLAASLSNAHNDVMDPCCTEHEVSVNRGSRVVMACNISNNLRDVTIELV 69
Qy 70 AHQGESAIENEVAPGYFSRDGQWLOVQGGVAQLVIKGARDSHAGLYMMHLVGHORNNRQV 129
Db 70 TSEKTSIIFNHTPPGNYKSDSWOLHIQGVQAOQLVITDAQDKHSGNYSWKLHGFQAEKFN 129
Qy 130 TLEVSGAEPOSAPDTGFWPVP-----AVVTAVFILLVALVFWAYRCRCSSQ 176


```

; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; TYPE: PRT
; LENGTH: 969
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (848)..(889)
; OTHER INFORMATION: Where Xaa is any amino acid
US-10-052-648A-35

Query Match          7.5%; Score 98; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 1.8;
Matches 51; Conservative 24; Mismatches 86; Indels 44; Gaps 10;

Qy      2 QTCPLAFPG---HVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSWGENT 51
Db      620 QACPPGFWGACFACSHCHN-----GASCSAEDGACHCTPGWTGLFCTQCPAFAFFGKDC 674
Qy      52 VMSCNISNAPS--HVNIIKLRAH-----GQESAIENEVAPGYFSRDGWOL-----QV 95
Db      675 GRVQCQNGASCDHISGKCTCRTGTGTQHCE--ORCAPGTGYGCGQQLCECMNNSTCDHV 732
Qy      96 QGG-VAQLVIKGARDSHAGLYMHVGHORNNRQVTLVSGAEPQSDPTGFWPVPVAVT 154
Db      733 TGTCTCSPGFGIRCDQALMMEELNPNYTKISPAL-----GAERHSV-----GAVTGIML 782
Qy      155 AVFILLVALVMFAWYRCRCQORRE 179
Db      783 LLFLIVLLGLFAWHRRRQKEGRD 807

RESULT 10
US-10-052-648A-34
; Sequence 34, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.

; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; TYPE: PRT
; LENGTH: 969
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (848)..(889)
; OTHER INFORMATION: Where Xaa is any amino acid
US-10-052-648A-34

Query Match          7.3%; Score 96; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 2.9;
Matches 51; Conservative 23; Mismatches 87; Indels 44; Gaps 10;

Qy      2 QTCPLAFPG---HVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSWGENT 51
Db      620 QACPPGFWGACFACSHCHN-----GASCSAEDGACHCTPGWTGLFCTQCPAFAFFGKDC 674
Qy      52 VMSCNISNAPS--HVNIIKLRAH-----GQESAIENEVAPGYFSRDGWOL-----QV 95
Db      675 GRVQCQNGASCDHISGKCTCRTGTGTQHCE--ORCAPGTGYGCGQQLCECMNNSTCDHV 732
Qy      96 QGG-VAQLVIKGARDSHAGLYMHVGHORNNRQVTLVSGAEPQSDPTGFWPVPVAVT 154
Db      733 TGTCTCSPGFGIRCDQALMMEELNPNYTKISPAL-----GAERHSV-----GAVTGIML 782
Qy      155 AVFILLVALVMFAWYRCRCQORRE 179
Db      783 LLFFIVLLGLFAWHRRRQKEGRD 807

RESULT 11
US-10-072-012-446
; Sequence 446, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
```



```

; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 446
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-072-012-446

Query Match
Best Local Similarity 7.3%; Score 95.5; DB 15; Length 506;
Matches 51; Conservative 25; Mismatches 79; Indels 37; Gaps 8;

QY 4 CPLAPGHVSQALGTLFLAASLSAQNKGWDSPICTEGVSVSVSGENTVMSCNISNAP-S 62
DB 228 CEVA---HVTLGGPPLRGNTANLS---ETIRVPPTLEITRSPSAGNQVNVTCQVNFYPR 281
QY 63 HVNKLRAHG-----QESAIINEVAPGFESRDGQLOVQGGVAQLVIKIGARDSHAGLYMW 117
DB 282 HLQUTWLENGMNSRTEAASVLVENKDGTFNOTSMLLVNNSAHRRAVLTQ----- 332
QY 118 HLVGHQHNR---NRQVTLSVGAEPQAPDTGTFWPVP-----AVVTAVFILLVALVM 165
DB 333 ---VEHDROPVSKNHTLEVSA--PQKDDQDTGTGCPNDNNNTSIFIVGVVVCALLVALLI 388
QY 166 FAWYRCRCSQQR 177
DB 389 AALYLLRIRONK 400

RESULT 12
US-10-719-993-534
; Sequence 534, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF

```

```

; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 534
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-993-534

Query Match
Best Local Similarity 7.0%; Score 91.5; DB 16; Length 541;
Matches 51; Conservative 39; Mismatches 89; Indels 77; Gaps 11;

QY 11 HVQSALGTLFLAASLSAQNKG-----WDGPICTEGVSVSVSGENTVMSCNISNAP-S 62
DB 155 HLQVYGSLLLIALLSSYRKGRGDEGFWKQVQYQTPRRSPFRSTISWEQV-----SPYSAYL 210
QY 63 HVNKLRAHQESAIINEVAPGFESRDGWL-----QVQGGVAQLVIKIGARDSHA 112
DB 211 HDAVLLVAETVKQV---KAGGDF-QDGWOLVSALKGSSQTTVGGITGTPVFVDAQGERHM 266
QY 113 G-----LYMMHLVGHQNRNNRQVTLVSGAEPQAPDTGFW-----PVPVAV 153
DB 267 DYSVYALQKSENGPLLLSLFLHYDSYQRNFNMNTPW-HGSLPDRPGCGFYNELCETQSAF 325
QY 154 TAVFILLVALVMFAWYRCRCSQQRREKKFFLEPOMKVAALRAGAQQGLSRASAEWL--- 210
DB 326 TGTAVLLTLM-----ILIPVGAAIL--GLILRMQRQNKDIWQI 364
QY 211 -----TPDSEPTPR 219
DB 365 NPDITILPONKPSQR 380

RESULT 13
US-10-425-114-62807
; Sequence 62807, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62807
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-208-G7_FLI.pep
US-10-425-114-62807

Query Match
Best Local Similarity 7.0%; Score 91.5; DB 15; Length 597;
Matches 62; Conservative 44; Mismatches 84; Indels 121; Gaps 16;

QY 21 FLAASLSAQNKGWDS-----ICTGVSVSVSW-----GE 49
DB 299 FXAAAEAGHPDWELPDDAGEINDTPTDGTGFTTAERGTYLTQGRFFLTWYSKRLIQHGD 358
QY 50 NTVMSCNISNAPSHVNIKLRA-----HQESAINEVAPGFES-----RDGWQLQVQ 96
DB 359 RVLDEAN--KAFGLCKVKLAAKVSGIHWYRHPSHAA---ELTSGYINLGRDGY----- 408

```



```
Db 1779 HGAQVLDIAIY---SCRVGAEQDFPVQVEEVAAKFCRLLEPVCGLGGTVTTLACELSPA 1835
Qy 61 FSHVNIKLRAHQESAI FNEVAPGVFSRDGHQLOVQGGVAOLVIKIGARDSHAGLYMMHLV 120
Db 1836 CAEVVWRC-----GNTQPRVGKR--FQVVAEGPVRSLTVLGLRAEDAGEY-----V 1879
Qy 121 GHQNNR---QVTLEV-----SGAEPQSAPDTGFMFVPVPAVTVAVFILLVALYMFAYR 170
Db 1880 CESRDDHTSAQLTVSVPRWKFMSGLSTVVAEEGGEATFQCWSPSDVAVV-----WFR 1933
Qy 171 CRCQORREKFFLEPOMKVAALRAGAQGLSRASAEJ-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSEKFAISQSGASHSLTISDLVLEDAGQITVEAEGASSAALRV 1983
Qy 226 KPSPLGALLELLSPQ 239
Db 1984 REAPVLFKKKLEPQ 1997
```

Search completed: October 17, 2005, 07:56:58
Job time : 2068 secs

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	1308	100.0	248	2	US-08-755-559-1	Sequence 1, Appli	
2	1308	100.0	248	3	US-09-210-474-1	Sequence 1, Appli	
3	1308	100.0	248	3	US-09-539-774-1	Sequence 1, Appli	
4	1308	100.0	248	4	US-09-997-165-4	Sequence 4, Appli	
5	1308	100.0	248	4	US-09-949-016-6397	Sequence 6237, Ap	
6	1308	100.0	258	4	US-09-949-016-8807	Sequence 8807, Ap	
7	326	24.9	212	4	US-09-997-165-8	Sequence 8, Appli	
8	90.5	6.9	171	4	US-09-489-033A-10298	Sequence 10298, A	
9	84.5	6.5	381	4	US-09-252-991A-27771	Sequence 27771, A	
10	84.5	6.5	455	4	US-09-949-016-6949	Sequence 6949, Ap	
11	84.5	6.5	455	4	US-09-949-016-11026	Sequence 11026, A	
12	83.5	6.4	199	3	US-09-430-503-46	Sequence 46, Appli	
13	83.5	6.4	199	3	US-09-430-503-48	Sequence 48, Appli	
14	83.5	6.4	1512	4	US-09-949-016-10490	Sequence 10490, A	
15	83	6.3	152	4	US-09-270-767-43788	Sequence 43788, A	
16	82.5	6.3	288	1	US-08-396-650-1	Sequence 1, Appli	
17	82.5	6.3	288	1	US-08-768-626-1	Sequence 1, Appli	
18	82.5	6.3	288	4	US-09-645-069-12	Sequence 12, Appli	
19	81.5	6.2	199	3	US-09-430-503-42	Sequence 42, Appli	
20	81.5	6.2	199	3	US-09-430-503-44	Sequence 44, Appli	
21	81.5	6.2	209	3	US-09-430-503-18	Sequence 18, Appli	
22	81.5	6.2	209	3	US-09-430-503-20	Sequence 20, Appli	
23	81.5	6.2	269	3	US-09-430-503-2	Sequence 2, Appli	
24	81.5	6.2	269	3	US-09-430-503-4	Sequence 4, Appli	
25	81.5	6.2	269	3	US-09-430-503-6	Sequence 6, Appli	
26	81.5	6.2	269	3	US-09-430-503-8	Sequence 8, Appli	
27	81.5	6.2	310	4	US-09-328-353-5485	Sequence 5485, Ap	

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDSPTPRPLALVFKPSPLGALELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDSPTPRPLALVFKPSPLGALELLSPQ 240
QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 2

US-09-210-474-1
; Sequence 1, Application US/09210474
; Patent No. 6072034
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-210-474-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDSPTPRPLALVFKPSPLGALELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDSPTPRPLALVFKPSPLGALELLSPQ 240
QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 3

US-09-539-774-1
; Sequence 1, Application US/09539774
; Patent No. 6350615
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/539,774
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,474
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-539-774-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWQLQVGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCSCQORREK 180

Qy 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQQORREK 180
Db 131 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQQORREK 190
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKSPGALBLLSPQ 240
Db 191 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKSPGALBLLSPQ 250
Qy 241 LFPYAADP 248
Db 251 LFPYAADP 258

RESULT 7

US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997.165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-997-165-8

Query Match 24.9%; Score 326; DB 4; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.9e-28;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
Qy 10 GHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNAFHVNIKLR 69
Db 10 GLFPRMLWALLLAASLNAHNDVWDEPCCCTHEVSVNRGSRVWACNISNLRDVTIELV 69
Qy 70 AHGQESAIFNEVAPGYFSRDGKQVQGVQVLAQVKGARDSHAGLYMMHLVGHQNNRQV 129
Db 70 TSEKTSIIFNHTPGNYKSDSKQLHIQGVQVLAQVITDAQDKHSGNSYKSLHGFQAEKFN 129
Qy 130 TLEVSGAEPQSDPTGFWPVP-----AVTVAFILLVALVMFAWYRCRCQQ 176
Db 130 NLTVNAADRKQKTDLPVTKVDPKPPTAVRTEVIIIAIATITITIGVFWY----- 182
Qy 177 RREKKFFLLEPQMKVA 192
Db 183 ----KQFPVAPQIOMS 194

RESULT 8

US-09-489-039A-10298
; Sequence 10298, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10298
; LENGTH: 1171

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10298

Query Match 6.9%; Score 90.5; DB 4; Length 1171;
Best Local Similarity 22.7%; Pred. No. 1.1;
Matches 57; Conservative 37; Mismatches 96; Indels 61; Gaps 13;
Qy 19 LLFLAASLSAQNEGWDSPICTEGVSV-----VSMGENTV---MSCNISNAFHV 64
Db 582 LAFYAASL-ASHPOWELPV-DDGLVSOVTRTLRIQLQORNSESTLYQKMLAQVANQYADM 639
Qy 65 NI-KLRAHGOESAIP--NEVAPGYFSRDGKQVQGVQVLAQVKGARDSHAGLYMMHLVG 121
Db 640 RLADMTADTASRLFTSTDEVVPGMFTQAWEQVQVPAIEKVAE-RRDE----MDWVLSD 694
Qy 122 HORNRRQVT-----LEVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQQ 177
Db 695 TKOTAAQSTSPALRALRYFADFSGAW-----LDFLNSLRWQRAATLSDA 742
Qy 178 REKKFFLLEPQMK--VAALRAGAQOQ-----LSRASAEIWTDPDSEP----- 216
Db 743 IDQLTLMADVRSPLVALMNTLSVQGRGTGOTGEAIDSLVKSARQLFNDRNSPVIDORSG 802
Qy 217 TRPLALVFKP 227
Db 803 ARGPLDATEFP 813

RESULT 9

US-09-252-991A-27771
; Sequence 27771, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27771
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27771

Query Match 6.5%; Score 84.5; DB 4; Length 381;
Best Local Similarity 26.2%; Pred. No. 0.91;
Matches 60; Conservative 27; Mismatches 91; Indels 51; Gaps 12;
Qy 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 192 MVTAPLMPEVITGLSLLLFVA---MAQLIGWPQ---ERGI VTI-WIAHT-----SFC 238
Qy 61 FSHVNIKLRAGHGOESAIFNEVAPGYFSRDGQW---LQVQGVQVLAQVKGARDSHAGLYMW 117
Db 239 SSYVAVVVSARURELDSIEEAMD LGAKPKWFLITITPMIAPSAAGGMSFA-LSLD 297
Qy 118 HLVGHQRNNRQVTVLEYSV-----AEPQSDPTGFWPVPVAVTAVFILLVAL-VNFAWY 169
Db 298 DLV-----LASFVSGPGSTTLPMEVFSVRLGVKPEINAVASLILLSVSLFTFFAWY 349
Qy 170 RCRCSQORREKKFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTP 218
Db 350 FTRQAEERR-----RRAIQQAEMEETATD-WQKGS-PTP 380

RESULT 10


```
US-09-949-016-6949
; Sequence 6949, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6949
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6949

Query Match          6.5%; Score 84.5; DB 4; Length 455;
Best Local Similarity 22.2%; Pred. No. 1.2;
Matches 41; Conservative 33; Mismatches 58; Indels 53; Gaps 10;

QY      8  FPGH-VSQALGTLFLAASLAQNE-----GWDSPICTEGVVSVSGENTVMSNCISNAFS 62
      4  FPGYNLSGAVASFLFILLTMKQSEDFRVIGPAHPILA-GV-----GEDALITCOLLPKRT 57
      63  HVNIKLRAHGOESAI FNEVAPGYFSRD-----GWQLQVOGGVAQ-----LVIK 105
      58  TMHVEVRYRSEPS-----TPVFVHRDGVETEMQMEEYRCGWVWEIENGIAKGNVALKIH 112
      106  GARDSHAGLYMHLV-CHQRRNRQVTLVSGAEPSQAPD-----TGFW 147
      113  NIQPSDNGQYWCHFDQNGYCGTSLLLKVAGL--GSAPSIHMEGPGSGVOLVCTARGWF 170
      148  PVPVAV 152
      171  PEPQV 175

RESULT 11
US-09-949-016-11026
; Sequence 11026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11026
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11026

Query Match          6.5%; Score 84.5; DB 4; Length 455;
Best Local Similarity 22.2%; Pred. No. 1.2;
Matches 41; Conservative 33; Mismatches 58; Indels 53; Gaps 10;

QY      8  FPGH-VSQALGTLFLAASLAQNE-----GWDSPICTEGVVSVSGENTVMSNCISNAFS 62
      4  FPGYNLSGAVASFLFILLTMKQSEDFRVIGPAHPILA-GV-----GEDALITCOLLPKRT 57
      63  HVNIKLRAHGOESAI FNEVAPGYFSRD-----GWQLQVOGGVAQ-----LVIK 105
      58  TMHVEVRYRSEPS-----TPVFVHRDGVETEMQMEEYRCGWVWEIENGIAKGNVALKIH 112
      106  GARDSHAGLYMHLV-CHQRRNRQVTLVSGAEPSQAPD-----TGFW 147
      113  NIQPSDNGQYWCHFDQNGYCGTSLLLKVAGL--GSAPSIHMEGPGSGVOLVCTARGWF 170
      148  PVPVAV 152
      171  PEPQV 175

RESULT 12
US-09-430-503-46
; Sequence 46, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-503-46

Query Match          6.4%; Score 83.5; DB 3; Length 199;
Best Local Similarity 21.9%; Pred. No. 0.44;
Matches 34; Conservative 28; Mismatches 54; Indels 39; Gaps 6;

QY      39  TEGVSVSVW-----GENTVMSNCISNAFSHVNIKLRAGHGOESAI FNEVAPGYFSRDGWQL 93
      66  TGLTSLVSVMSFQPEGADTTVS-----FPHYS-----OGQVYLGNYPPFKDRISWAG 111
      94  QVQGGVAQLVIKGARDSHAGLYMHL-----VGHQRRNRQVTLVSGAEPSQAPDT 144
      112  DLKDKASINENMQFIHNGTYICDVKNPPDIVVQPGH-----IRIYVVEKENLPVP 163
      145  GFWPVPAVTVAFI---LLVALYMFAYRCRCQQ 176
      164  PVMWVAVTVAVVLGLTLILSMILAVLYRRKNSKR 198

RESULT 13
US-09-430-503-48
; Sequence 48, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-503-48
```



```

Query Match          6.4%; Score 83.5; DB 3; Length 199;
Best Local Similarity 21.9%; Pred. No. 0.44;
Matches 34; Conservative 28; Mismatches 54; Indels 39; Gaps 6;

Qy 39 TEGVVVSW-----CENTVMSCNISAFSHVNIKLRAHQESAIFNEVAPGVFSRDGWL 93
Db 66 TGLTTSVSWSQPEGADITVS-----PFHYS-----QQVIGNYPPFKDRISWAG 111

Qy 94 QVQGVAGLVITKGARDSHAGLYMMHL-----VGHQRNNRQVTLVSGAEPOSAPT 144
Db 112 DLDDKKDASINENMQFIHNGTVICDVKNPPDIVVQPGH-----IRYVVEKENLPVF 163

Qy 145 GFWPVPVAVTAVFI---LLVALVMEFAWYRCRCSEQ 176
Db 164 PWWVVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

RESULT 14
US-09-949-016-10490
; Sequence 10490, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10490
; LENGTH: 1912
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10490

```

			6.4%; Score 83.5; DB 4; Length 1912;	
			Best Local Similarity 20.7%; Pred. No. 14;	
			Matches 62; Conservative 31; Mismatches 90; Indels 117; Gaps 15;	
Query Match				
Best Local Similarity	6.4%;	Score 83.5;	DB 4;	Length 1912;
Matches	62;	Conservative 31;	Mismatches 90;	Indels 117; Gaps 15;
QY	5	PLAFIG--HYVSOALGTLFLLAASLSAQNEGH-----DSPTCTEGVVSVSW--	47	
		: : : : : : : : : : : : : : : : :		
Dd	592	PLNPPGDTHITVPL-ALMFEDVTVAATNFSPDYCSAVQALEAAAPCRACVGSIWRCHWCP	650	
QY	48	-----CENTVMSCNISNAFSHVNLIKRAHGQESAIINEVAPGYFSRDGHQLQ	94	
		: : : : : : : : : : : : : : : : :		
Dd	651	QSSHCVVGEHCPEGETIYSA-----QEVDIQVRPGCACPQVEGLAGP-HLVIPVGWESH	703	
QY	95	VQGGAVALVIKGARDSHAGLWMHLGVHORNNRQVTTLEVSGAEPOSAPDTGF-----	146	
	:	: : : : : : : : : : : : : : : : :		
Dd	704	LALVRNL--QHFRGLPASFCWLFLPGLRLGLPATLE-----ETAGDSGLIHCOAHQF	755	
QY	147	-----WPVPAYVT-----AVFTLLVALVMFAWYRCR-----CQQORREKKF	182	
		: : : : : : : : : : : : : : : : :		
Dd	756	YPSMSQRLELPVIYVTQGEARQLDNTHALYVL-----YDCAMGHDPDCHS-----	800	
QY	183	FLLEPMKVMVALRAGAOGGLSRASAELWTPDSEPTPR--PLUALVFKSPSLGABELLSPQ	240	
Dd	801	-----COANRSLGCLWCADGPACRYGPLC-----PPGAVELLPCAP	838	

```
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43788
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43788

Query Match      6.3%; Score 83; DB 4; Length 152;
Best Local Similarity 33.3%; Pred.No. 0.33;
Matches 31; Conservative 9; Mismatches 27; Indels 26; Gaps 7;

Qy   170 RRCSCSQREKFFLL---EPQWK-----VAALRAGAQOGLSRASAEIWTPDSEPTP- 218
           ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |
Db    51 RINCAPPREAVHHRIRDAHPQMGNRRRWLAPlQIGNQ--SR-----KQFI PP 99
           ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |

Qy   219 --RP-LALVFKPSPLGALELLSPQLFPYAADP 248
           ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |
Db   100 RRPALLAILAPSPFGSGQ-IAPSLLPVAPSP 131
           ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
```